

*John M. Stewart (Heidelberg)*

**Reading Comprehension of L2 Medical Texts:  
Steps, Structures, and Strategies**



**Editors:**

**Viktorija Bilić**

**Anja Holderbaum**

**Anne Kimmes**

**Joachim Kornelius**

**John Stewart**

**Christoph Stoll**

**Publisher:**

**Wissenschaftlicher Verlag Trier**

John M. Stewart (Heidelberg)

## Reading Comprehension of L2 Medical Texts:

### Steps, Structures, and Strategies<sup>1</sup>

#### Abstract:

This paper combines ideas from the literature on natural language parsing at the clause/sentence level with studies of reading comprehension at the larger discourse level in order to arrive at a holistic view of the psycholinguistic and cognitive processes involved in comprehending written LSP texts in the field of medicine. The already somewhat "foreign" nature of specialized discourse is compounded when the texts are written in the reader's L2, as is often the case in fields such as medical translation or foreign language instruction for the (medical) professions. The present contribution attempts to synthesize findings from theoretical and applied linguistics in order to recommend teaching and learning strategies for improved reading comprehension of L2 medical texts.

#### Keywords:

reading comprehension; language for special purposes; medical translation; sentence parsing; discourse comprehension; superstructures; LSP reading strategies

#### Contents:

1	The Reception Phase in LSP Translation .....	2
2	Sentence-Level Parsing .....	7
3	Text Comprehension .....	11
4	Strategies for Improving Comprehension .....	21
5	Conclusions .....	22
6	Bibliography .....	23

---

<sup>1</sup> This is a revised version of a paper presented at the *Teaching Medical Translation II (TMT II)* conference in Heidelberg, Germany on July 7, 2009 ([w<sup>1</sup>](#)).

## 1 The Reception Phase in LSP Translation

Reading written texts is a complex mental activity requiring the coordination of visual, attentional, linguistic, logical, and pragmatic/communicative faculties. This is true even of the most accessible everyday texts written in the reader's native language and in a familiar, informal register. The process becomes considerably more complex when the text in question deals with specialized fields of knowledge and human activity. Scientific and technical texts are usually written in so-called "languages for special purposes" (LSPs),<sup>2</sup> defined by Baakes (1994: 1) as follows: "A complete set of linguistic phenomena comprising terminological, syntactic, and stylistic features which are different from ordinary language and occur within a definite sphere of communication." Such texts pose special challenges for readers with varying degrees of expertise in the field concerned. These challenges can consist of both unfamiliar terminology and esoteric insider knowledge that is either tacitly presupposed as a prerequisite for comprehending the text or, in some cases, novel ideas and discoveries being presented for the first time by the text at hand.

Readers are even confronted with such challenges when the text is written in their native language (L1). In a manner of speaking, these L1 texts can attain a degree of "foreignness" comparable to that of a text written in the reader's second language (L2). The L1 reader is then placed in a similar situation to that of an advanced foreign language learner, who has mastered the basic structures and much of the lexicon of the L2, but must nevertheless overcome certain knowledge gaps in order to successfully construct a mental representation of the text. But of course the difficulties are further compounded when the LSP text is written in a second or foreign language. The text can then be conceived of as doubly foreign to the reader, since his or her lack of expertise in the specialist subject matter is combined with imperfect knowledge of the language of communication. A common scenario where this latter situation obtains is in university-level instruction in languages for the professions, where students of medicine, law, business, or other fields learn to use the corresponding LSP in a foreign language. A second example, and the one which concerns the present

---

<sup>2</sup> Also known as "technical language" or "technolects".

contribution, is the case of language professionals who translate LSP texts from some L2 into their L1. Because translators typically have a strong background including professional training in their L2, as a group they can be regarded as "advanced learners" in terms of the way they take on an L2 reading comprehension task. Thus, certain parallels can be drawn between LSP reading comprehension in the L1 and general reading comprehension in the L2. Since Nida/Taber (1969: 31), the translation process has often been viewed as comprising three phases: reception, transfer, and formulation. The first of these three is the stage in which a translator reads the source text, activates background knowledge, and takes steps such as consulting reference works in order to achieve a sufficient comprehension of the text. During the second, or "transfer" phase, the translator begins to seek equivalent target-language expressions and to make strategic decisions (Hervey, Loughridge and Higgins 2006: 6), such as considering the degree to which the text will need to be localized, or adapted to the target culture. During the third, or "formulation" phase, the translator proceeds to actually produce the target-language text. Most of the activities involved in comprehending an L2 source text fall into the reception phase, which will therefore be the focus of attention in the present paper.

But the question arises, "Do translators really need to fully comprehend a source text?" After all, lay observers often see translation as merely repeating someone's words in the other language. Contrary to this widespread misconception, however, we can answer the question with a resounding "Yes!". A prerequisite for competently translating any text is to have comprehended it at a level comparable to that of the average native speaker. In the translation of everyday texts, this is necessary in order to avoid pitfalls such as false friends, to recognize and properly convey irony, culture-specific attitudes, and implicit assumptions, among other things. When it comes to translating LSP texts, the need for full comprehension becomes even more urgent due to the extensive background knowledge that readers of such texts are generally presupposed to have. To name just one example, a German source text dealing with the anatomy and function of the heart speaks of *Spannung* in a particular context. This word is polysemous; one sense translates into English as "tension," and another as "(electric) charge." Regardless of whether the translator's native language is

German or English, there is a potential for misinterpretation here, because on the one hand muscular tension is a common topic of discussion in texts about the heart, and on the other hand, so is electric current. In this case the word *Spannung* occurred in the following context:

A) Mit Hilfe von EKG-Elektroden, eines Verstärkers und eines Schreibers können die Veränderungen der elektrischen **Spannung** an der Körperoberfläche gemessen und aufgezeichnet werden. Diese **Spannung** hat zu jedem Zeitpunkt eines Herzschlags eine bestimmte Ausrichtung und Stärke. (w<sup>2</sup>)

Translation: *With the help of EKG electrodes, an amplifier, and a recorder, the changes in the electric **charge** on the surface of the body can be measured and recorded. At any given point during a heartbeat, this **charge** has a certain orientation and magnitude.*

This can be contrasted with the following excerpt from an introductory text on the anatomy and function of the human heart:

B) Das Herz wird überwiegend in der Diastole mit Blut versorgt. In der Systole steht die Herzwand wegen der Muskelkontraktion unter **Spannung** und lässt kaum Durchblutung zu. (w<sup>3</sup>)

Translation: *The heart is supplied with blood mainly during diastole. During systole, the heart wall is under **tension** due to the muscle contraction and scarcely permits any blood flow.*

In excerpt A), the German term *Spannung* clearly can only be translated in the electrical sense (i.e. "charge"), since the text is explaining how the heart's conduction system (a set of specialized muscle cells that conduct electric signals to the myocardial muscle tissue of the heart, thus stimulating it to contract, pumping the blood out of a given chamber) generates a certain voltage that is detectable even on the outside of the chest. In excerpt B), however, a different kind of *Spannung* is being referred to; here it is the tensile force that occurs when muscles contract, or "tension." A translator who lacks sufficient medical background knowledge may fail to recognize in which sense *Spannung* is being used, especially if confronted with a text that provides fewer contextual clues than the two cited above. The

choice between the two possible English equivalents "tension" and "charge" must therefore be informed by an understanding of the biological processes at work here. In the highly demanding field of medical conference interpreting, it is not uncommon for interpreters to continue beyond their undergraduate studies, even pursuing doctoral degrees in medicine in order to master the subject matter necessary for understanding spoken medical professional discourse (Stoll 2009).

Kjær (2000) argues along similar lines concerning the importance of background knowledge for understanding legal LSP texts. She sets up a typology of legal texts comprising 1) the descriptive type, which "describe" legal rules and serve the purpose of enlarging the reader's knowledge of such rules; 2) the constitutive type, which "lay down" legal rules and serve the reader's purpose of identifying pertinent rules or formulating rules; and 3) the reproductive type, which "apply" legal rules and serve the purpose of identifying rules, for example in order to enable the reader to follow rule-based argumentation. Descriptive texts, such as legal textbooks, are the only type that is fully explicit with regard to the legal rules involved. Reproductive texts, such as legal judgments, often refer to the legal conditions and effects of a rule without stating its "connecting concepts." Constitutive texts such as statutes – though from an a priori point of view they would seem necessarily to be just as explicit as descriptive texts – serve as more of a basis from which the rules can be extracted and interpreted in legal proceedings than a statement of the rules themselves. It follows, according to Kjær, that reproductive and constitutive texts, due in part to their high degree of intertextuality, will not be fully accessible to the uninitiated reader. This is enough to preclude translators without legal expertise from competently translating such texts.

Aside from these difficulties of LSP texts, which arguably are inevitable, or a "necessary evil" resulting from the knowledge gradients between the authors and readers of texts, readers often encounter "unnecessary" difficulties as well. LSP texts constitute part of the broader phenomenon of specialist communication, which Baumann defines as follows:

... die (Fach)Kommunikation [ist] als gedanklich-sprachliche Interaktion von Kommunikationspartnern zu verstehen, in deren Verlauf fachbezogene Informationen ausgetauscht und (mehr oder weniger effizient) verarbeitet werden. (1996: 95)

Translation: ... *(specialist) communication [is] to be understood as a cognitive-linguistic interaction, in the course of which field-specific information is exchanged and (more or less efficiently) processed.*

The key phrase here may be "more or less efficiently," since indeed, LSP texts are not always composed optimally for the sake of comprehension. Consider the following sentence excerpted from the abstract of a scholarly contribution to an edited volume:

This chapter argues that claims for recurrent networks as plausible models of human sentence processing from which generalizations that have seemed to require the mediation of symbolically represented grammars are "emergent" are misplaced. (Steedman 2002: 95)

We cannot rule out the possibility that the author deliberately formulated this sentence in a convoluted fashion (including complex noun phrases and multiple embeddings), as a sort of inside joke among specialists in natural language processing. Presumably the sentence is to be parsed so as to render the following syntactic dependencies:

This chapter argues that claims<sub>x</sub> for recurrent networks<sub>y</sub> as plausible models of human sentence processing<sub>y</sub> from which generalizations<sub>z</sub> that have seemed to require the mediation of symbolically represented grammars<sub>z</sub> are "emergent<sub>z</sub>" are misplaced<sub>x</sub>.

In other words, it is the "claims" that are misplaced, the "recurrent networks" that are posited as plausible models of human sentence processing, and the "generalizations" that are emergent and seem to require mediation. Assuming my interpretation of this excerpt is

correct, then in order to render the sentence more readily comprehensible – so that the reader is not compelled to reread it several times – we might rephrase it as follows:

Some researchers have claimed that recurrent networks represent a plausible model of human sentence processing. They further claim that certain generalizations are "emergent" from these networks, seeming to require the mediation of symbolically represented grammars. This chapter argues that such claims are misplaced.

Of course, the author may simply have been motivated by considerations of economy; our more explicit paraphrase is 25% longer than the original. But although the specialist communication stands to gain a kind of efficiency (in Baumann's terms) through brevity, it may suffer a net loss if readers have to struggle to parse the sentence. This of course applies to translators as well as to the L1 audience. Such "unnecessary" difficulties combine with the "necessary" ones to create a formidable obstacle to comprehension on the part of translators.

In this paper, I will attempt to briefly characterize the nature of the psycholinguistic task of comprehending complex written discourse, in a foreign language, within the field of medicine. German and English will serve as the language pair for examples and illustrations, but the argumentation to be employed here should be applicable across most languages in which such specialist discourse is written. After establishing a theoretical framework based on models of sentence parsing on the one hand and text/discourse comprehension on the other, I will apply insights gained within these fields to specific excerpts from medical texts. This will be followed by some conclusions regarding the most effective learning and teaching strategies that can be practiced in order to improve the L2 medical reading comprehension of translators and other concerned professionals.

## **2 Sentence-Level Parsing**

To proceed from the smallest units to the largest, we must first go below the level of the sentence to that of the individual word. Since LSPs are known primarily for their esoteric

terminology, what part do individual lexical items play in the overall comprehensibility of LSP texts? Hoffmann sees the lexical component of LSPs as their primary distinguishing feature:

The specificity of LSPs manifests itself especially clearly in their **vocabularies**, where each of them has created its own more or less autonomous **terminology**, which has become a subsystem of the overall **lexical** system of the language in question. In the *grammar*, there is no technical-language subdomain or subsystem. What we find are rather restrictions in the usage of syntactic and morphological means, while the basic normative rules are followed. (1998: 190) [emphasis mine]

In the case of everyday or general texts, Nation (2006) finds that readers need to know 98% of the lexical items used in a written text for unassisted reading comprehension. This corresponds to a passive vocabulary of 8,000 – 9,000 word families.<sup>3</sup> Laufer and Sim (1985) carried out a similar investigation with regard to academic texts in the English language, finding that a score of 65-70% or better was needed on the First Certificate of English exam to enable readers to comprehend the specialized type of texts. They were even able to rank various factors according to their importance for a correct interpretation of the text's meaning: most important was vocabulary, followed by knowledge of the subject, and finally the clues that can be drawn from discourse markers and syntax. The professional literature thus seems to corroborate many observers' intuitive impression that it is the terminology that makes LSPs so difficult to understand<sup>4</sup>. Similar conclusions were reported by Stewart (2008): medical English is characterized not so much by its own distinct syntax, but by a preponderance of certain stylistic devices such as nominal groups, the passive voice and metonymy (which serve purposes of impersonalization), and markers of epistemic modality used in order to hedge and express uncertainty.

---

<sup>3</sup> A "word family" can be defined as a stem and all the forms transparently derived from it.

<sup>4</sup> On the other hand, Levine and Haus (1985) found that background knowledge could be more important than language level in comprehending Spanish L2 texts about baseball. But it should be borne in mind that "language level" here refers to general language proficiency, and not to familiarity with the specialized terminology.

But how does a reader proceed from the lexical level to the phrasal, clausal, and sentence level, and are there any major differences between this process when applied to written as opposed to spoken language? Consider the following sentence: "It frightened the patient that the physician wanted to perform surgery." In spoken language, the listener is confined to the same temporal one-dimensionality as the speaker; that is, the sentence must be uttered – and perceived – phoneme by phoneme, word by word, etc., with little overlap between the individual units in the incoming speech signal. This may constitute the chief difference between the parsing of written versus spoken language, since written language can often be simultaneously perceived at several different levels: word, sentence, paragraph, page, and even the text as a whole. It follows that a reader has more clues at his or her disposal for parsing a sentence than a listener has. This should help the reader to more quickly eliminate those candidate parsings which will ultimately prove to be incorrect for the sentence at hand. In other words, listeners should be more susceptible to straying down "garden paths" than readers.<sup>5</sup> Both readers and hearers of this sentence need to create a mental representation of its syntax like that depicted (in simplified form) in Figure 1 below:

---

<sup>5</sup> "Garden paths" are constructions marked by some degree of syntactic ambiguity, which may not be resolved until the parser has already assigned (or begun to entertain) an incorrect interpretation. A classic example is the sentence "The horse raced past the barn fell," in which the word "raced" can either be (incorrectly) interpreted as a past-tense verb in the active voice or as a past participle in the passive voice. Filip et al. (2002) make a clever play on this reduced relative construction in the title of their article, "Reduced Relatives Judged Hard Require Constraint-Based Analyses."

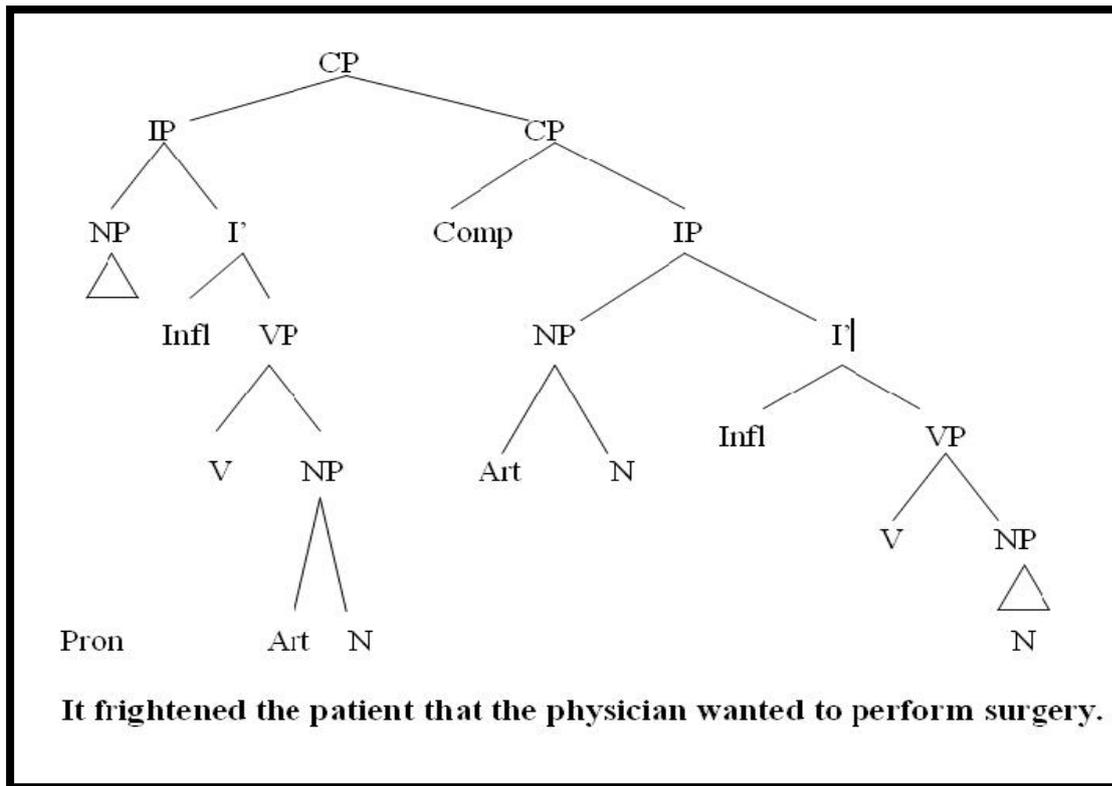


Fig. 1: Tree diagram of syntactic structures generated during sentence parsing

This sentence could give rise to a garden path insofar as a listener may initially interpret the word "it" as a third-person neuter pronoun instead of an expletive subject. But in both the written and the spoken media, the syntactic processor may be biased against this erroneous interpretation by the preceding discourse (if, for example, no non-human entity has been mentioned which could serve as an antecedent of the pronoun). At a later point in the sentence, the word "that" could trigger a different garden path if it is interpreted in the function of a relative pronoun coindexed with the noun phrase "the patient." Only upon reaching the end of the sentence can the listener be certain that the sentence will not take the form "It frightened the patient that the physician wanted to perform surgery on when the lights went out." However, it can be argued that a reader will be less inclined to entertain such garden-path parses than a listener, if only because peripheral vision (and attention) reveals that the sentence will end with the word "surgery." On the other hand, the written medium may also present the parser with certain disadvantages due to this forward-looking dimension; when the eyes "jump ahead" to a later point in the sentence

without having sufficiently parsed the first few words and phrases, garbled readings can ensue. But I would argue that such phenomena are not peculiar to medical LSP texts.

Since the syntax of medical texts seldom diverges significantly from the general norms for written language, the only aspect of written medical LSP which could present special difficulties in terms of sentence-level parsing would be at the lexical/morphological level. If, for example, many nouns within the terminology were homonymous with verbs, adjectives, etc., then the reader might stumble while attempting to assign the correct morphosyntactic category to such terms. The non-LSP lexicon contains numerous homonyms of this kind, as in the sentence "He saw her duck" (discussed in Crocker and Corley 2002: 162), or the above case of "that" as subordinating conjunction vs. "that" as relative pronoun, or "it" as 3<sup>rd</sup> person neuter pronoun vs. "it" as expletive subject (see Figure 1). But in fact, medical terminology often renders the distinction between parts of speech more systematic and explicit rather than less. For example, the verb "anesthetize" is clearly differentiated from the noun "anesthesia" or the adjective "anesthetic." The only homonymous pair in this word family would then be the adjective "anesthetic" and the noun whose spelling and phonetic realization are identical to it (e.g., "an anesthetic agent" as opposed to "a local anesthetic"). The net effect of these considerations is that readers' comprehension of medical texts is most likely to be hindered not so much *syntactically* at the level of sentence parsing, but at other levels. The problem of assigning a semantic interpretation to unfamiliar lexical items has been briefly sketched above. The other key source of difficulty – comprehension at the text or discourse level – will be addressed in the next section.

### 3 Text Comprehension

The historical roots of theories of discourse comprehension date back to classical antiquity, to the ancient Greek philosophers' thoughts concerning rhetoric. Their contributions were often concerned with the process of text production (ways of structuring arguments, persuasively making a case, etc.) more than with understanding and interpreting texts. But there is an unbroken tradition, from antiquity through the Middle Ages to the present, of hermeneutics and exegesis. These are primarily directed at enabling readers to interpret

literary, allegorical, and prophetic texts, with their multiple layers of meaning and sophisticated use of symbolism. More recently – i.e. since the 20<sup>th</sup> century – linguists, pedagogues, and psychologists have become interested in how the human mind processes spoken and written language in general, based on a contemporary understanding of natural language grammar and neural function. Psycholinguists in particular have built up complex theoretical models taking into account the need for a phoneme-detection system, a word-recognition device, a syntactic processor, and a semantic/discourse processor (Gee 1993). This model is broad enough in scope to span from the individual phonemes or letters of a text, to its phrases and clauses, to the discourse as a whole. It provides an account of how the mind processes these different units and constituents to arrive at a mental representation of the sentence, text, or discourse in question. The structure of words, phrases, clauses, and sentences has been addressed by morphologists and syntacticians in the structuralist and generative traditions, as illustrated in Figure 1. Semanticians have represented the compositional meaning of entire sentences with logical operators and formalisms such as the lambda calculus. What remained to be accounted for was the meaning-structure of entire texts and discourses.

Enter van Dijk and Kintsch, with their seminal 1983 book *Strategies of Discourse Comprehension*. In this work, the authors describe both the hierarchical information structures that constitute the meanings of whole texts, as well as the mental faculties, models, and representations that are needed in order to understand those meanings. They argue that mental representations of texts and discourses have five key properties: they are 1) *constructivist* (meaning that representations are not absorbed in some ready-made form, but the mind must play an active role in constructing them); 2) *interpretative* (i.e., the mental representation consists not only of the perceived visual or linguistic data, but also an interpretation of their significance); 3) *online* (understanding takes place progressively as the input is coming in, not post-hoc once the end of the discourse has been reached); 4) *presuppositional* (understanding involves activating background knowledge, beliefs, motivations, etc.); and 5) *strategic* (the understander seeks to construct a mental representation as effectively and efficiently as possible). The most crucial of these five for

our current purposes is the strategic manner in which such mental representations are formed, since strategies can be learned, taught, and differentially employed by applied linguists, such as language teachers and translators.

Another important part of the model developed by van Dijk and Kintsch (1983) is the division of memory capacities between the working (or short-term) memory and episodic memory. Working memory plays a key role, in that individual readers and listeners differ in terms of the amount of data they can hold in this short-term storage. And, as van Dijk and Kintsch point out, the familiar constant of  $7 \pm 2$  items that can be maintained in short-term memory (Miller 1956) only applies to simple data such as a list of numbers. When the items not only require storage, but also some degree of manipulation (establishing logical relations between them, etc.) the working memory capacity drops to as few as two or three items. Yet van Dijk and Kintsch cite research showing that it is not the general short-term memory capacity that differs between good readers and poor readers, but the so-called "reading span" (memory span within a reading task) (1983: 352).

Since even the short-term memory of individuals with above-average memory capacities is insufficient for storing all of the data and logical relationships that make up a short text, readers and listeners must have recourse to other forms of memory. Long-term memory serves to store whatever knowledge may be permanently retained from a text, but for online processing during the actual reading or listening task, it is useful to be able to maintain a number of partially processed sets of data (often referred to as "chunks") to be reaccessed and further manipulated as the text progresses. This is where episodic memory comes in; here, information and inferences that have been gleaned from the text can be kept in a sort of "holding pattern" in case they are needed for constructing further representations of the discourse. By way of analogy, the individual bits of data a reader or listener encounters in a text can be thought of as balls that he or she is juggling. A moderately proficient juggler can usually manage three balls at a time. But if the performer's assistant begins to throw in additional balls, the juggler can quickly become overwhelmed. Episodic memory functions like a shelf at shoulder's height, where the juggler

(reader/listener) can set aside individual balls in order to lighten the processing load, while keeping them within reach for later use.

Alongside the developing representation of this "text base," van Dijk and Kintsch also highlight the ongoing process of updating and activating a *situation model*, or "the cognitive representation of the events, actions, persons, and in general the situation, a text is about" (1983: 12). This situation model is also familiar from studies of sentence-level parsing and comprehension. Gee illustrates it with the example of two sentences: "1) Three turtles rested on a floating log and a fish swam beneath them. 2) Three turtles rested on a floating log and a fish swam beneath it" (1993: 273). Though there are lexical and semantic differences between these two sentences, experimental subjects generally cannot correctly identify which of these sentences they have heard a short time before. This is because both sentences result in the same situation model in the hearer's mental representation, and this model is all that is stored in memory. Additional useful notions for understanding how the mind represents situations include such cognitive-science concepts as schemas, frames, scenarios, and scripts. These refer to different ways in which the mind is thought to schematically represent the elements of a situation model (including the arguments or parties to an event or action, the component parts of a multi-step process, etc.).

Against this backdrop of mental faculties for understanding discourse, van Dijk and Kintsch also subdivide the strategies employed by listeners and readers into five main types: *propositional strategies*, which assign propositions to sentential surface structures (at a ratio of roughly one proposition per clause); *local coherence strategies*, which establish meaningful connections between successive propositions of a discourse; *macrostrategies* for drawing inferences about macropropositions (which in turn can be strung together over several levels to form the macrostructure, or gist, of the text); *schematic strategies* for recognizing the superstructure, or schematic structure of the text; and *production strategies* for the planning and composing of discourse (1983: 13-19). Van Dijk and Kintsch cite the following as examples of the sorts of logical relations that can obtain between successive propositions: "You've done a great job. I'll pay you double." "It's cold in here. Can you please shut the window?" (1983: 86). Though not made explicit, the second sentence in each pair

follows as a consequence of the first. Similar logical relations exist in LSP texts, though readers may require more training in order to recognize them.

The superstructure varies, of course, with the text type or genre. For the *story* genre, for example, van Dijk and Kintsch propose three main components that make up the "narrative schema": setting, complication, and resolution. Since this is one of the areas where the most extensive conventions have been developed, it has been targeted as an especially profitable object of study in translation and the other language professions. When readers are familiarized with the conventional structure of a certain genre, they have a head start in terms of recognizing the schematic structure of any future texts they may encounter which belong to that genre. This can then be used as a scaffolding aid in order to begin deducing certain propositional and macropropositional contents of the text which might otherwise have eluded the reader or listener.

As mentioned above, the visual aspect of written texts not only makes them markedly different from spoken texts in terms of how they are parsed at the sentence level, but even at the level of the entire text or discourse. This is because a single glimpse of the written text can often suffice in order to provide the reader with a holistic or *gestalt* impression, which in turn sets the presuppositional component of the discourse comprehension process in motion by activating background knowledge about text conventions and superstructures. Figure 2 illustrates this with two examples of texts shown at a greatly reduced size. Most lay readers will immediately recognize the genre of the text on the left-hand side. Perhaps only medical professionals – or those involved with the medical field in some other capacity – will be able to identify the genre of the text on the right-hand side:

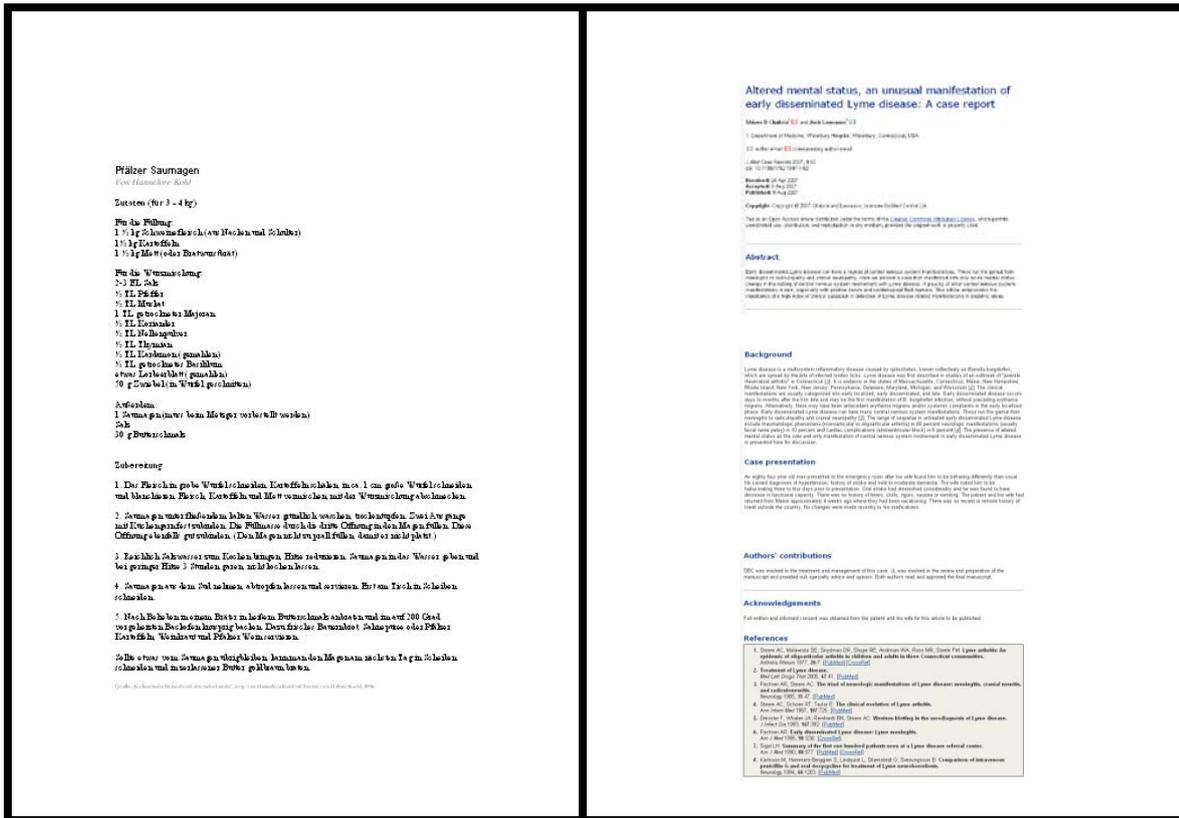


Fig. 2: Gestalt impressions of two text types recognizable from their macrostructure alone

The text on the left is a recipe, which can be deduced from the superstructure elements of 1) the list of ingredients and 2) the cooking instructions, arranged into a series of steps. The text on the right is an article from a medical journal, as revealed by its superstructure comprising a title, author’s names and affiliations, an abstract, the expository text, and the list of references.

This strategy of focusing on superstructures is precisely the approach advocated by Reitbauer (2000). In a study of scientific abstracts from professional journals in the field of psychology, Reitbauer argues that such abstracts are structured according to conventional schemes, and that (L2) readers can improve their reading comprehension of this essential text type by familiarizing themselves with the typical schematic structure, or superstructure. According to Reitbauer, the four main components of a typical scientific abstract are the topic, method, results, and conclusions, as represented by the diagram in Figure 3. By using graphic representations of superstructures as an instructional tool, Reitbauer found that

students' reading comprehension of the L2 scientific abstracts was positively influenced compared to a control group who received no such training with superstructures as orientation aids.

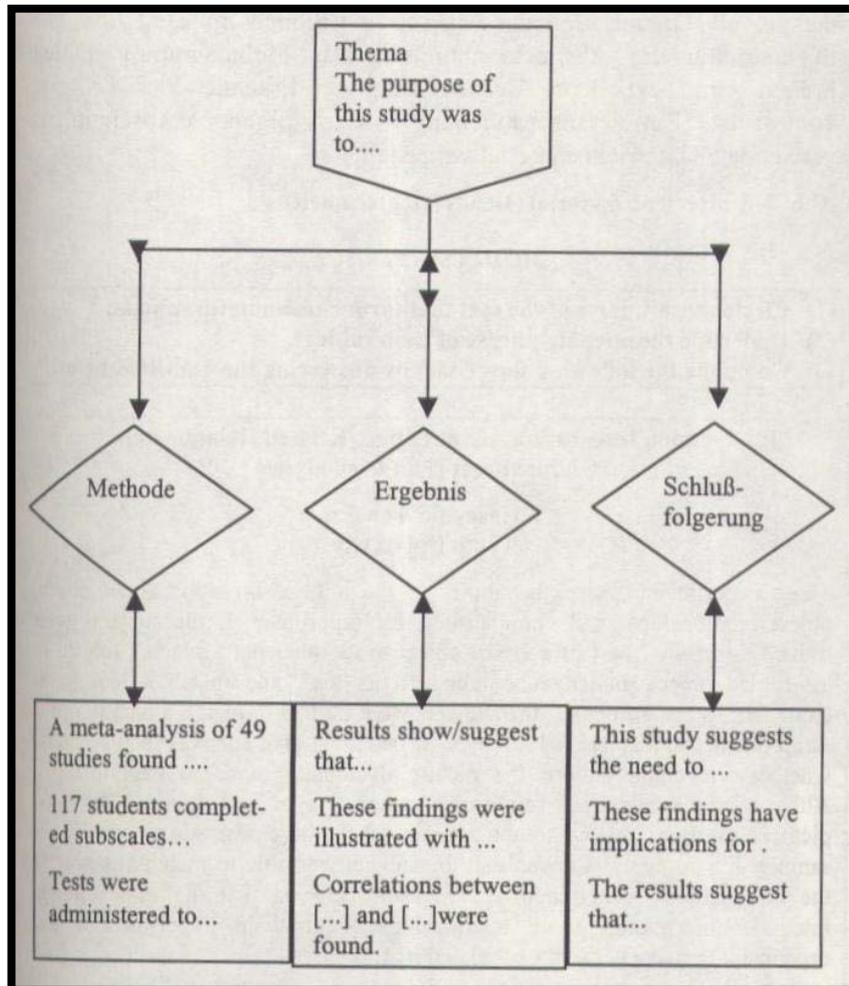


Fig. 3: The superstructure of psychology abstracts (from Reitbauer 2000: 270)

As a way to implement these structures for instructional purposes, Reitbauer suggests using flowchart-like diagrams with certain elements strategically omitted, which students are then required to fill in. This develops skills such as recognizing logical premises and conclusions,<sup>6</sup> assigning elements of the text to the results, conclusions, or other components of the

<sup>6</sup> Especially important in this context are discourse markers, which students can learn to highlight in order to identify the pivotal points in the argumentation.

superstructure, and identifying the key findings and gist of the scientific text. Figure 4 shows one of the examples with which Reitbauer illustrates this teaching method with regard to abstracts. Similar tools could be employed with a wide array of (longer) medical text genres: journal articles, pharmaceutical package inserts, summaries of product characteristics, case reports and medical histories, referral letters, press releases and other public relations texts, among others.

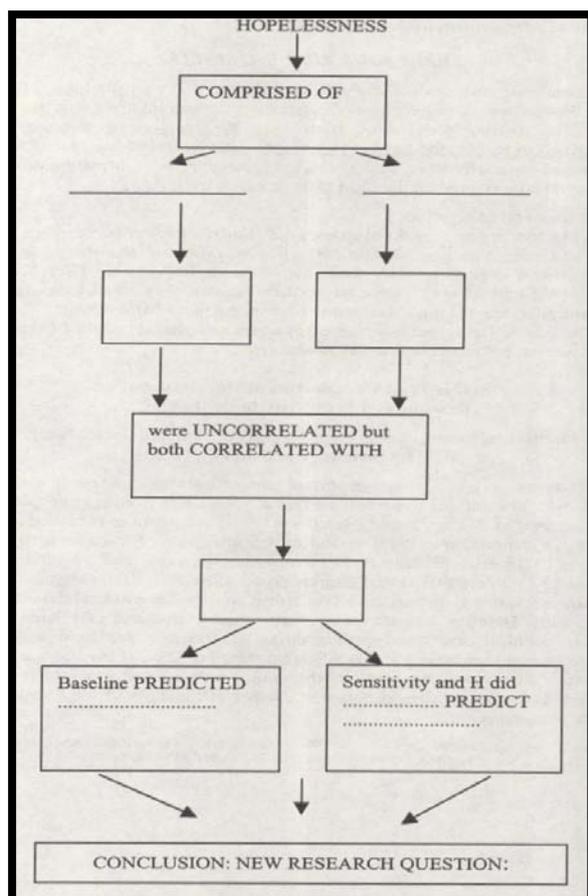


Fig. 4: A fill-in-the-blank superstructure flowchart (from Reitbauer 2000: 172)

This approach could prove quite useful for developing students' competence in dealing with the macrostructure of texts. But as argued above, another crucial component of LSP comprehension (in both L1 and L2) is the individual lexeme, the term itself. Figures were cited from the professional literature with regard to the "threshold" for comprehension, meaning the minimum percentage of lexical items used in a text which the reader must be

familiar with in order to understand that text. If we test these claims against a short medical text (see Figure 5, the "Summary" section of a medical journal article), we find that the specialized terms do indeed constitute the main obstacle to comprehension by lay readers. My German-L1 students in an undergraduate medical technical translation course reported that the terms appearing in red print below were unknown to them:

**Medical Research Council randomised trial of endometrial resection versus hysterectomy in management of menorrhagia**

**Background**  
 The most frequent indication for hysterectomy is menorrhagia, even though the uterus is normal in a large number of patients. Transcervical resection of the endometrium (TCRE) is a less drastic alternative, but success rates have varied and menorrhagia can recur. We have tested the hypothesis that the difference in the proportion of women dissatisfied and requiring further surgery within 3 years of TCRE or hysterectomy would be no more than 15%.

**Methods**  
 202 women with symptomatic menorrhagia were recruited to a multicentre, randomised, controlled trial to compare the two interventions. TCRE and hysterectomy were randomly assigned in a ratio of two to one. The primary endpoints were women's satisfaction and need for further surgery. The patients' psychological and social states were monitored before surgery, then annually with a questionnaire. Analysis was by intention to treat.

**Findings**  
 Data were available for 172 women (56 hysterectomy, 116 TCRE); 26 withdrew before surgery and four were lost to follow-up. Satisfaction scores were higher for hysterectomy than for TCRE throughout follow-up (median 2 years), but the differences were not significant (at 3 years 27 [96%] of 28 in hysterectomy group vs. 46 [85%] of 54 in TCRE group were satisfied;  $p=0.053$ ). TCRE had the benefits of shorter operating time, fewer complications, and faster rates of recovery.

**Interpretation**  
 TCRE is an acceptable alternative to hysterectomy in the treatment of menorrhagia for many women with no other serious disorders.

Fig. 5: A medical LSP text with the unfamiliar medical terms highlighted

(Text taken from [Glendinning and Holmström 1998: 61](#))

Interestingly, the unfamiliar specialized terms amount to 27 word tokens, or 11% of the 248 words in the text – well above the 2% threshold allowed by [Nation \(2006\)](#) – but there is a great deal of lexical recurrence, so that these terms comprise only seven different "types":

*endometrial, resection, hysterectomy, menorrhagia, indication, uterus, and transcervical.*<sup>7</sup>

This puts the unfamiliar terms in a different light, numerically speaking; a small investment of effort toward learning these seven types will enable most readers to comprehend the entire 248-word text.

The learning task can be made even more efficient by breaking some of these terms down into their component parts, through morphological analysis. *Endometrial* is made up of *end-* 'inside' + Gk. *mētra* 'uterus,' and thus refers to the inside lining of the uterus. *Hysterectomy* breaks down into Gk. *hystera* 'womb' + *-ectomy* 'cutting out,' hence: the surgical removal of the uterus. *Menorrhagia* consists of Gk. *menos-* 'month (gen.)' + *rhage* 'breaking/bursting,' and means abnormally profuse menstrual bleeding. *Transcervical* is formed from the prefix *trans-* 'across' and the Latin stem *cervix* 'neck' (here the narrowing of the uterus) and refers to a technique for accessing the uterus via the cervix in order to remove (*resection*) the endometrium. Knowing a fairly small number of Greek and Latin roots can thus open the door for students to morphologically analyze a vast number of medical terms. In the present example, the translator or other lay reader can use this lexical knowledge to "bootstrap" his or her way to understanding the macrostructure of the text (with the help of the section-marking provided by the typological formatting). Ultimately, he or she arrives at an understanding that this medical journal article is presenting research comparing two different procedures for treating excessive menstrual bleeding: hysterectomy, which entails complete removal of the uterus, versus the less drastic alternative of transcervical resection of the endometrium (TCRE), which removes the endometrium but leaves the uterus in place. The results suggest that, since there was no statistically significant difference in satisfaction rates between the two procedures, TCRE is an acceptable alternative in the absence of other serious disorders. This example illustrates, then, how non-specialist L2 readers of medical LSP texts can be trained to take on reading comprehension challenges from two directions:

---

<sup>7</sup> That is, if we count members of a single "word family" (e.g. "endometrium" and "endometrial") as one and the same type. Also, "TCRE" has been treated here not as a separate term, but as a transparent composite of "transcervical resection of the endometrium", since it is indicated as such in the text.

bottom-up by way of morphological analysis and lexical familiarization,<sup>8</sup> and top-down by way of local coherence strategies, macrostrategies, and schematic strategies.

#### 4 Strategies for Improving Comprehension

Returning to my claim from the first part of this article, that LSP reading bears similarities to L2 reading, we can draw a number of helpful strategies for the learning and teaching of LSP reading skills from the extensive professional literature in L2 reading. When confronted with unfamiliar terms, students should be encouraged to first attempt morphological analysis. If this does not succeed, they can resort to contextual guessing (deducing the meaning of the unfamiliar term by inferring the meaning of the sentence as a whole, and then decomposing the compositional semantics to arrive at a meaning for the term in question). This process can be aided by so-called "fuzzy parsing", or assigning interpretations that include some margin of ambiguity. Tolerance of ambiguity or uncertainty in general is an indispensable asset when it comes to building up listening or reading proficiency in a new linguistic code. Beginning and intermediate readers of medical LSP texts must therefore learn to set realistic goals and to see partial comprehension as a successful outcome until full comprehension becomes feasible. In order to foster this attitude, teachers do not need to manipulate the text materials they present to their students so as to shorten them or render them more easily comprehensible. Rather, as practiced in L2 reading courses over the past decade, the key is to simplify the task, not the text. A useful exercise might then be to read a medical journal article (written *by* specialists *for* specialists) without looking up most of the unfamiliar technical terms, and with an eye toward understanding the main gist rather than every nuance of scientific reasoning and methodology.

---

<sup>8</sup> The phrase "lexical familiarization", by the way, has been used to refer to specific techniques employed by authors in order to help readers gain familiarity with less well-known terms, for example by providing a paraphrase or gloss in parentheses. Bramki and Williams (1984) recommend training students to recognize and make use of these comprehension aids.

## 5 Conclusions

To summarize the main points of this paper, the primary obstacles to (L2) non-specialist readers' comprehension of medical LSP texts do not appear to include the parsing of sentence-level syntactic constructions; rather, it is the lexis on the one hand and the macrostructure of the information contents and logical relations on the other. Many insights from foreign language learning and cognitive/psycholinguistics can be applied in order to boost learners' proficiency in reading such texts. Educators should focus their students' efforts on:

- Building morphological analysis skills
- Recognizing discourse markers
- Learning genre-specific superstructures

By doing so, they can help "bridge the gap"<sup>9</sup> between medical expert authors and non-specialist readers.

---

<sup>9</sup> An allusion to the motto of the 2009 TMT II conference: *Bridging the Gap: Medical Curriculum Meets Translation Methodology* ([w<sup>1</sup>](#)).

## 6 Bibliography

- Baakes, Klaus (1994). *Key Issues of Syntax in the Special Languages of Science and Technology*. Heidelberg: Groos.
- Baumann, Klaus-Dieter (1996). "Die Mehrdimensionalität des Fachtextverstehensprozesses. Fachtextsorten und Kognition". Gerhard Budin (ed.). *Multilingualism in specialist communication*. Proceedings of the 10th European LSP Symposium. Vienna Aug. 29 - Sept. 1, 1995. Vienna: International Network for Terminology. 95-112.
- Bramki, Doudja and Ray Williams (1984). "Lexical Familiarization in Economics Text, and its Pedagogical Implications in Reading Comprehension". *Reading in a Foreign Language* 2, 1. 169-181.
- Crain, Stephen and Mark Steedman (1985). "On not being led up the garden path: the use of context by the psychological syntax processor: The use of context by the psychological syntax processor". David R. Dowty Lauri Karttunen, and Arnold Zwicky (eds.). *Natural Language Parsing: Psychological, computational, and theoretical perspectives*. Cambridge University Press. 320-358.
- Crocker, Matthew and Steffan Corley (2002). "Modular Architectures and Statistical Mechanisms: The case from lexical category disambiguation". Paola Merlo and Suzanne Stevenson (eds.). *The lexical basis of sentence processing: Formal, computational and experimental issues*. Amsterdam: Benjamins. 157-180.
- Filip, Hana et al. (2002). "Reduced Relatives Judged Hard Require Constraint-Based Analyses". Paola Merlo and Suzanne Stevenson (eds.). *The lexical basis of sentence processing: Formal, computational and experimental issues*. Amsterdam: Benjamins. 255-279.
- Gee, James P. (1993). *An Introduction to Human Language: Fundamental concepts in linguistics*. Englewood Cliffs, NJ: Prentice-Hall.
- Glendinning, Eric and Beverly Holmström (1998). *English in Medicine: A course in communication skills*. 2<sup>nd</sup> ed. Cambridge University Press.

- Grütz, Doris, Harald Pfaff, and Jürgen Belgrad (2007). "Rezeptionsstrategien zum Verstehen von Sachtexten: eine Studie in Haupt- und Realschule Klasse 7". *Zeitschrift für angewandte Linguistik* 46. 57-83.
- Hervey, Sándor, Michael Loughridge and Ian Higgins (2006). *Thinking German translation: a course in translation method, German to English. 2<sup>nd</sup> ed.* London: Routledge.
- Hoffmann, Lothar (1998). "Fachsprachen als Subsprachen". Lothar Hoffmann et al. (eds.), *Fachsprachen: Languages for special purposes. An International Handbook of Special-Language and Terminology Research. Vol. 1.* Berlin, New York: de Gruyter. 189-199.
- Kendeou, Panayiota and Paul van den Broek (2007). "The effects of prior knowledge and text structure on comprehension processes during reading of scientific texts". *Memory & Cognition* 35, 7. 1567-1577.
- Kjær, Anne Lise (2000). "On the structure of legal knowledge: The importance of knowing legal rules for understanding legal texts". Lita Lundquist and Robert J. Jarvella (eds.). *Language, text, and knowledge: Mental models of expert communication.* Berlin; New York: Mouton de Gruyter. 127-161.
- Laufer, Batia and Donald D. Sim (1985). "Measuring and Explaining the Reading Threshold Needed for English for Academic Purposes Texts". *Foreign Language Annals*, 18:5 (Oct.). 405.
- Levine, Martin G. and George J. Haus (1985). "The Effect of Background Knowledge on the Reading Comprehension of Second Language Learners". *Foreign Language Annals*, 18:5 (Oct.). 391.
- Lundquist, Lita (1996). "Interpreting LSP texts. Background knowledge and scalar inferences". Gerhard Budin (ed.). *Multilingualism in specialist communication.* Proceedings of the 10th European LSP Symposium. Aug. 29 - Sept. 1, 1995. Vienna: International Network for Terminology. 239-258.

- Lundquist, Lita (2000). "Knowledge, events, and anaphors in texts for specific purposes". Lita Lundquist and Robert J. Jarvella (eds.). *Language, text, and knowledge: Mental models of expert communication*. Berlin; New York: Mouton de Gruyter. 97-126.
- Miller, George A. (1956). "The magical number seven, plus or minus two: Some limits on our capacity for processing information". *Psychological Review* 63 (2): 81-97.
- Millis, Keith K. and Arthur C. Grasseur (1994). "The time-course of constructing knowledge-based inferences for scientific texts". *Journal of memory and language* 33, 5. 583-599.
- Nation, Paul (2006). "How large a vocabulary is needed for reading and listening?". *Canadian modern language review* 63, 1. 59-82.
- Nida, Eugene Albert and Charles R. Taber. *Theorie und Praxis des Übersetzens* (German translation of: *The Theory and Practice of Translation*). Weltbund der Bibelgesellschaften, 1969.
- Reitbauer, Margit (2000). *Effektiver lesen mit Superstrukturen: Eine empirische Untersuchung anhand der Fachtextsorte Abstract*. Tübingen: Narr.
- Steedman, Mark (2002). "Connectionist and symbolist sentence processing". Paola Merlo and Suzanne Stevenson (eds.). *The lexical basis of sentence processing: Formal, computational and experimental issues*. Amsterdam: Benjamins. 95-108.
- Stewart, John M. (2008). "Syntax Patterns in Medical Texts". Presentation at the *Teaching Medical Translation* conference, Heidelberg, on July 3, 2008.
- Stoll, Christoph (2009). *Jenseits simultanfähiger Terminologiesysteme – Methoden der Vorverlagerung und Fixierung von Kognition im Arbeitsablauf professioneller Konferenzdolmetscher*. Heidelberger Studien zur Übersetzungswissenschaft 13. Trier: Wissenschaftlicher Verlag Trier.
- Tartter, Vivien C. (1998). *Language and its normal processing*. Thousand Oaks, Calif.: Sage Publications.
- van Dijk, Teun and Walter Kintsch (1983). *Strategies of Discourse Comprehension*. San Diego: Academic Press, Inc.

## Online Sources

- w<sup>1</sup> University of Heidelberg, Seminar for Translation and Interpreting (2009). Teaching Medical Translation Conference 2009. <http://www.tmt-heidelberg.com/en/konferenz.html>.
- w<sup>2</sup> Uniqa (2009). Homepage of the Austrian Uniqa online health portal. <http://www.meduniqa.at/622.0.html>.
- w<sup>3</sup> Kleindienst, Ralf (2009). Online electrocardiography handbook authored by a nurse specializing in intensive care and anesthesiology. <http://www.ekg-online.de/anatomie/anatomie2.htm>.

### **T21N - Translation in Transition**

T21N offers a cutting-edge electronic publishing venue, created by experts for both young talent and established researchers from the worlds of translation and interpreting.

T21N provides a stage for emerging ideas and new academic talent to present their ideas in a digital reading site, where speed and ease meet enjoyment.

T21N is exclusively published online at <http://www.t21n.com>.

Articles in compliance with our style sheet may be submitted at any time and will be published at short notice.

T21N editors research and teach at the Institute of Translation and Interpreting at the University of Heidelberg in Germany.

Editors: Dipl.-Übers. Viktorija Bilić, Dr. Anja Holderbaum, Dr. Anne Kimmes, Prof. Dr. Joachim Kornelius, Dr. John Stewart, Dr. Christoph Stoll